

PCT/JP98/04772

日 本 国 特 許 庁

PATENT OFFICE
JAPANESE GOVERNMENT

21.10.98

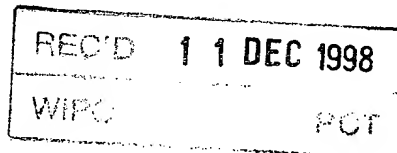
S

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application:

1997年10月22日



出 願 番 号
Application Number:

平成 9年特許願第289982号

出 願 人
Applicant(s):

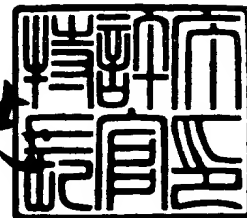
株式会社ヘリックス研究所

PRIORITY DOCUMENT

1998年11月27日

特許庁長官
Commissioner,
Patent Office

伴佐山 建志



出証番号 出証特平10-3094484

【書類名】 特許願

【整理番号】 H1-806

【提出日】 平成 9年10月22日

【あて先】 特許庁長官 殿

【国際特許分類】 C12N 15/00

【発明の名称】 完全長 cDNA クローンの選択方法

【請求項の数】 7

【発明者】

 【住所又は居所】 千葉県木更津市請西 2-16-13-401

 【氏名】 太田 紀夫

【発明者】

 【住所又は居所】 千葉県木更津市貝渕 3-9-17-509

 【氏名】 西川 哲夫

【発明者】

 【住所又は居所】 千葉県木更津市請西 2-3-13

 【氏名】 サラモフ アサフ

【発明者】

 【住所又は居所】 千葉県木更津市貝渕 3-9-17-606

 【氏名】 磯貝 隆夫

【特許出願人】

 【識別番号】 597059742

 【氏名又は名称】 株式会社ヘリックス研究所

 【代表者】 野口 照久

【代理人】

 【識別番号】 100102978

 【弁理士】

 【氏名又は名称】 清水 初志

【選任した代理人】

 【識別番号】 100108774

【弁理士】

【氏名又は名称】 橋本 一憲

【手数料の表示】

【予納台帳番号】 041092

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 要約書 1

【包括委任状番号】 9716404

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 完全長cDNAクローンの選択方法

【特許請求の範囲】

【請求項1】 完全長cDNAクローンを選択する方法であって、

(a) cDNAライブラリーに含まれるcDNAクローンの5'末端から塩基配列を決定する工程、

(b) (a)で決定された塩基配列中に開始コドンが含まれているか否かを、開始コドンを予測するプログラムによって判定する工程、

(c) (b)で開始コドンが含まれていると判定されたクローンを選択する工程、
を含む方法。

【請求項2】 cDNAライブラリーが完全長率の高いcDNAライブラリー作製法により作製されたものである、請求項1に記載の方法。

【請求項3】 cDNAライブラリーがmRNAのCAPを修飾する工程を含む方法により作製されたものである、請求項1に記載の方法。

【請求項4】 完全長cDNAライブラリーを作製する方法であって、

(a) cDNAライブラリーに含まれるcDNAクローンの5'末端から塩基配列を決定する工程、

(b) (a)で決定された塩基配列中に開始コドンが含まれているか否かを、開始コドンを予測するプログラムによって判定する工程、

(c) (b)で開始コドンが含まれていると判定されたクローンを選択する工程、

(d) (c)で選択されたクローンを混合する工程、
を含む方法。

【請求項5】 cDNAライブラリーが完全長率の高いcDNAライブラリー作製法により作製されたものである、請求項4に記載の方法。

【請求項6】 cDNAライブラリーがmRNAのCAPを修飾する工程を含む方法により作製されたものである、請求項4に記載の方法。

【請求項7】 請求項4に記載の方法により作製しうるcDNAライブラリー。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、遺伝子工学の分野に属し、完全長cDNAクローンを選択する方法に関する。

【0002】

【従来の技術】

現在、種々の動植物および微生物においてゲノムプロジェクトが進展し、多くの遺伝子が単離されその機能の解析が試みられているが、単離した遺伝子の機能を効率よく解析するためには完全なタンパク質を発現可能なcDNAクローン、即ち、完全長のcDNAクローンを効率よく得ることが重要なステップとなる。

【0003】

完全長率の高いcDNAライブラリーの作製法としては、現在のところ次のような方法が知られている。即ち、mRNAのCAPにRNAリンカーを酵素的に結合するオリゴキャップ法(菅野・丸山, 蛋白質 核酸 酵素, 38, 476-481 (1993).、鈴木・菅野, 蛋白質 核酸 酵素, 41, 603-607 (1996).、M. Maruyama and S. Sugano, Gene, 138, 171-174 (1994).)、オリゴキャップ法とOkayama-Berg法を組み合わせで開発した改良オリゴキャップ法(S. Kato et al., Gene, 150, 243-250 (1994).、加藤・関根 特開平6-153953号公報 1994年6月3日)、CAPにDNAリンカーを化学結合するリンカー化学結合法(N. Merenkova and D. M. Edwards, WO 96/34981 Nov. 7, 1996.)、CAPのビオチン修飾によるキャップ化学修飾法(P. Carninci et al., Genomics, 37, 327-336 (1996).、P. Carninci et al., DNA Research, 4, 61-66 (1997).)である。これらはすべて真核生物のmRNAのCAPを修飾する方法であり、完全長率の高いcDNAライブラリーを作製する方法である。また、Capをトラップすることにより完全長cDNAの比率の高いライブラリーを作製する方法としては、5' Capサイトの標識法として、酵母またはHela細胞のCap結合タンパク質を用いる方法(I. Edery et al., MCB, 15, 3363-3371 (1995))などが知られている。これら以外の方法としては、第一鎖cDNAの5' 端をC-テイリングしてCap Switch oligonucleotideをアニールさせるCap Switch oligonucleotide法であるCap Find

er (Clontech社製) が知られている。

【0004】

しかしながら、従来の方法に比して完全長率が高いとはいえ、これら方法を用いてcDNAライブラリーを作製した場合には、不完全長のクローンが必ずある頻度で混入してしまう。そこで、遺伝子の高効率機能解析を行い、新規な有用遺伝子を高効率でクローン化するために、cDNAライブラリーに含まれる各クローンが完全長であるか否かを容易に識別しうる方法の開発が望まれていた。

【0005】

【発明が解決しようとする課題】

本発明は、効率的に完全長のcDNAクローンを選択する方法、および完全長率の高いcDNAライブラリーの作製方法を提供することを課題とする。

【0006】

【課題を解決するための手段】

本発明者等は、上記課題を解決すべく鋭意研究を行った結果、5'末端の一定領域が欠失したcDNAには翻訳開始コドンが含まれない可能性が高く、完全長cDNAには翻訳開始コドンが存在することから、翻訳開始コドンの存在の有無を指標として、cDNAライブラリーから効率的に完全長cDNAを選択することができると考えた。特に、完全長率の高いcDNAライブラリー作成法で作製したcDNAライブラリーを用いれば、高効率で完全長cDNAを選択しうると考えた。即ち、完全長率の高いcDNAライブラリー作成法でcDNAライブラリーを作成後、5'末端より数百塩基のDNA塩基配列を決定し、この領域内に翻訳開始コドンが存在するか否かを解析して翻訳開始コドンが存在するクローンを選択することにより、効率的に完全長のcDNAクローンを単離することができると考えた。

【0007】

しかしながら、翻訳開始コドンを予測する方法に関し、cDNAの翻訳開始点を予測するプログラムに関しては、ほとんど開発がなされていないのが現状である（報告例としては、「A. G. Pedersen, Proceedings of fifth international conference on intelligent systems for molecular biology, p226-233 (1997), held in Halkidiki, Greece, June 21-26, 1997.」が挙げられる）。エキソンを

予測するプログラムについては開発されてきてはいるが（「Gene Finder」V. V. Solovyev et al., Nucleic Acids Res., 22, 5156-5163 (1994).、「Grail」Y. Xu et al., Genet-Eng-N-Y., 16, 241-253 (1994)）、このプログラムを用いることのみによって完全に翻訳開始点を決定することはできない。

【0008】

そこで、本発明者等は、独自にcDNAの翻訳開始コドンの予測ソフトを開発した。そして、完全長率の高いcDNAライブラリー作成法で作製したcDNAライブラリーに含まれるクローンの5'領域の塩基配列を決定し、この開発したソフトを利用してこの5'領域に翻訳開始コドンが存在するか否かの検討を行った。具体的には、まず、オリゴキャップ法で全長cDNAライブラリーを作製し、cDNAライブラリーに含まれるいくつかのクローンについてその5'末端側の塩基配列を決定し、決定した塩基配列に基づき、クローンをデータベース上で既知であるクローンと新規なクローンとに分別した。次いで、翻訳開始コドンの予測ソフトを用いて5'末端側の決定された塩基配列につき翻訳開始コドンの有無およびその存在位置を判定し、既知のクローンについては予測ソフトで認定された翻訳開始コドンと、データベース上の翻訳開始コドンの存在位置との整合性を検討した。その結果、本発明者等は、検討を行った既知のクローンにつき、予測ソフトにより判定された翻訳開始コドンの有無および存在位置が実際のデータベース上での事実と一致することを見いだした。

【0009】

ここに本発明者等は、独自に開発したソフトが翻訳開始コドンの存在の有無および存在位置を的確に判定することができ、このソフトにより翻訳開始コドンが存在すると判定されたクローンをcDNAライブラリーから選択すれば、効率的に完全長のcDNAクローンを選択することが可能であることを見いだした。さらに、本発明者等は、選択したクローンを混合することにより、完全長率の極めて高いcDNAライブラリーを作製できることを見いだした。

【0010】

即ち、本発明は、cDNAライブラリーから完全長cDNAクローンを選択する方法、およびこれにより選択したcDNAクローンを混合することを特徴とする完全長cDNA

ライブラリーの作製法に関し、より具体的には、

- (1) 完全長cDNAクローンを単離する方法であって、
 - (a) cDNAライブラリーに含まれるcDNAクローンの5'末端から塩基配列を決定する工程、
 - (b) (a)で決定された塩基配列中に開始コドンが含まれているか否かを、開始コドンを予測するプログラムによって判定する工程、
 - (c) (b)で開始コドンが含まれていると判定されたクローンを選択する工程、を含む方法、
- (2) cDNAライブラリーが完全長率の高いcDNAライブラリー作製法により作製されたものである、(1)に記載の方法、
- (3) cDNAライブラリーがmRNAのCAPを修飾する工程を含む方法により作製されたものである、(1)に記載の方法、
- (4) 完全長cDNAライブラリーを作製する方法であって、
 - (a) cDNAライブラリーに含まれるcDNAクローンの5'末端から塩基配列を決定する工程、
 - (b) (a)で決定された塩基配列中に開始コドンが含まれているか否かを、開始コドンを予測するプログラムによって判定する工程、
 - (c) (b)で開始コドンが含まれていると判定されたクローンを選択する工程、
 - (d) (c)で選択されたクローンを混合する工程、を含む方法、
- (5) cDNAライブラリーが完全長率の高いcDNAライブラリー作製法により作製されたものである、(4)に記載の方法、
- (6) cDNAライブラリーがmRNAのCAPを修飾する工程を含む方法により作製されたものである、(4)に記載の方法、
- (7) (4)に記載の方法により作製しうるcDNAライブラリー、に関する。

【0011】

【発明の実施の形態】

本発明は、cDNAライブラリー、特に完全長率の高いcDNAライブラリーの5'領域

の塩基配列を、翻訳開始コドンを高精度で予測するソフトにより解析することにより、効率的に全長cDNAクローンを単離することができ、さらに単離したcDNAクローンを混合することにより完全長率をさらに高めたcDNAライブラリーを作製することができるとの本発明者等による知見に基づく。従って、本発明の完全長cDNAクローンを選択する方法は、(a) cDNAライブラリーに含まれるcDNAクローンの5'末端から塩基配列を決定する工程、(b) 決定された塩基配列中に開始コドンが含まれているか否かを、開始コドンを予測するプログラムによって判定する工程、および(c) 開始コドンが含まれていると判定されたクローンを選択する工程を含む。また、本発明の完全長cDNAライブラリーの作製法は、さらに(d) 選択されたクローンを混合する工程を含む。

【0012】

本発明の方法において、5'末端領域の塩基配列の決定を行う「cDNAクローン」としては、特に制限はないが、クローンが完全長率の高くないライブラリーに由来する場合には、完全長率の高いライブラリーに由来する場合と比較して、結果として完全長cDNAの単離の効率が低下すると考えられる。このためcDNAクローンは、上記の完全長率の高いcDNAライブラリーの作成法、例えば、mRNAのCAPにRNAリンカーを酵素的に結合するオリゴキャップ法(菅野・丸山, 蛋白質 核酸 酵素, 38, 476-481 (1993).、鈴木・菅野, 蛋白質 核酸 酵素, 41, 603-607 (1996).、M. Maruyama and S. Sugano, Gene, 138, 171-174 (1994).)、オリゴキャップ法とOkayama-Berg法を組み合わせ開発した改良オリゴキャップ法(S. Kato et al., Gene, 150, 243-250 (1994).、加藤・関根 特開平6-153953 1994年6月3日)、CAPにDNAリンカーを化学結合するリンカー化学結合法(N. Merenkova and D. M. Edwards, WO 96/34981 Nov. 7, 1996.)、CAPのビオチン修飾によるキャップ化学修飾法(P. Carninci et al., Genomics, 37, 327-336 (1996).、P. Carninci et al., DNA Research, 4, 61-66 (1997).)、酵母またはHela細胞のCap結合タンパク質を用いる方法(I. Edery et al., MCB, 15, 3363-3371 (1995))、Cap Switch oligonucleotide法であるCap Finderにより作製されたライブラリーに由来することが好ましい。cDNAライブラリーからのcDNAクローンの単離は、文献(J. Sambrook, E. F. Fritsch & T. Maniatis, Molecular Cloning Second edition, Cold Spring

ng Harbor Laboratory Press 1989) などに記載の常法により行うことができる。

【0013】

クローンの5'末端からの塩基配列の決定は、例えば、アプライドバイオシステムズ社のDNAシーケンス試薬とDNAシーケンサー等を用いて常法により行うことができる。塩基配列は全配列を決定する必要はなく、5'末端から1000塩基程度を決定すれば十分であるが、500塩基程度、さらには300塩基程度の塩基配列の決定であっても高い精度が期待できる。

【0014】

クローンの5'末端からの塩基配列の解析に用いられる「開始コドン予測プログラム」としては、本発明者らにより開発された後述する実施例1に記載のプログラムが好ましい。決定された塩基配列中に開始コドンが含まれているか否かの判定は、プログラムによる解析の結果、抽出されるスコアにより行われる。スコアが高い値を示し決定した塩基配列中に開始コドンが存在すると判定されたcDNAクローンは、通常は、完全長cDNAであり、逆に、スコアが低い値を示し塩基配列中に開始コドンが存在しないと判定されたcDNAクローンは、不完全長cDNAである。従って、cDNAライブラリーから塩基配列中に開始コドンが存在すると判定されたcDNAクローンを選択することにより、効率的に完全長cDNAを単離することが可能である。実際に実施例1に記載のプログラムを用いた解析の一例では、完全長率が51%のcDNAライブラリーをクローンの選択に用いた場合、スコア（最高値0.94）を0.5以上でクローンを選択するとクローンの完全長率は71%、スコアを0.70以上でクローンを選択するとクローンの完全長率は77%、スコアを0.80以上でクローンを選択するとクローンの完全長率は81%、スコアを0.90以上でクローンを選択するとクローンの完全長率は85%であった。従って、実施例1に記載のプログラムを用いてスコアの高いクローンを選択することにより、高い確率で完全長cDNAクローンが選択される。

【0015】

また、本発明の完全長cDNAクローンの選択方法により選択されたクローンを混合することにより、クローンの選択に用いたcDNAライブラリーの完全長率を飛躍

的に高めたcDNAライブラリーを再構築することも可能である。このようなタンパク質に発現可能なcDNAからなるcDNAライブラリーをライブラリーのまま発現させれば、混合物としてタンパク質を発現させた高効率遺伝子機能解析系を構築することができ、これにより有用な遺伝子を高効率でクローニングすることが可能となる。

【0016】

以下、本発明を実施例によりさらに詳細に説明するが本発明はこれら実施例に制限されるものではない。

【0017】

【実施例】

【実施例1】 cDNAの翻訳開始コドン进行予測するプログラムの作製

本発明のcDNAの翻訳開始コドン进行予測するプログラムは、cDNA配列断片が与えられた時に、その中に含まれる全てのATGの中から真の翻訳開始コドンらしいものを予測する。予測は、基本的には、A) 着目したATGの両側の一定範囲領域（数10～数100塩基）のそれぞれにおける翻訳領域らしさの情報と、B) 着目したATGの付近の翻訳開始コドンらしさの情報をを用いて行う。あらかじめ、翻訳領域と非翻訳領域が判明している配列を多数用いて、翻訳領域と翻訳開始コドン付近の配列の特徴を抽出しておく。プログラムは、これらの特徴情報をを用いて翻訳開始コドン进行予測する。

【0018】

ゲノムエクソン予測プログラムであるGene Finder (Solovyev V.V., Salamov A.A., Lawrence C.B. Predicting internal exons by oligonucleotide composition and discriminant analysis of spliceable open reading frames. Nucleic Acids Res. (1994) 22: 5156-63.) 中で用いられている線形判別分析を、予測の最適化の手法として用いた。線形判別分析では、データの持つ複数の特徴情報をそれぞれ数量化し、それらに重みをつけて加算した量をスコアとして用いた。ここでは、スコアを翻訳開始コドンらしさの確率（翻訳開始コドンがすでに判明しているデータを用いた場合の正解率を確率とする）に変換する。即ち、与えられたcDNA配列断片に含まれる各ATGに対して翻訳開始コドンらしさの確率が出

力される。翻訳開始コドンの認定は、翻訳開始コドンらしさの確率が一定のしきい値を超えるかどうかで行うが、しきい値はその後の解析の方針に応じて、即ち、どの程度のノイズを許容して解析を行うかに応じて、その値を設定する。例えば、40%のノイズを許容可能であれば、しきい値0.6を採用すればよい。重みのパラメーターは、翻訳開始コドンがすでに判明しているデーターをトレーニングデーターとして用いて、予測制度が最大になるように決定する。使用した特徴情報は、上述のAとBの情報を具体化したそれぞれ3つの情報を用いた。

【0019】

具体的には、A) 着目したATGの両側の一定範囲領域（数10～数100塩基）のそれぞれにおける翻訳領域らしさの情報としては、①ATGからストップコドンまで（最大で、ATGの300塩基後まで）に含まれる6塩基文字の頻度情報、②ATGの前後50塩基に含まれる6塩基文字の頻度情報の差、③シグナルペプチドらしさの指標（ATGの後30アミノ酸（90塩基）の中で、最も疎水的な8アミノ酸文字の疎水性指標）を用いた。また、B) 着目したATGの付近の翻訳開始コドンらしさの情報としては、①ATGの14塩基前から5塩基後までの領域内の3塩基文字を単位とした重み行列情報、②ATGの前に同じフレームで他のATGがあるかないか（あれば1、なければ0）、③ATGの36塩基前から7塩基前までの領域に含まれるシトシン塩基の頻度を用いた。

【0020】

[実施例2] オリゴキャップ法によるcDNAの調製、および翻訳開始コドン予測プログラム（ATGpr）による解析

オリゴキャップ法でcDNAライブラリーを作成し、個々のクローンについて常法によりプラスミドDNAを抽出した。具体的には、ヒト胎盤およびヒト培養細胞（teratocarcinoma NT-2、neuroblastoma SK-N-MC）より文献（J.Sambrook, E.F.Fritsch & T.Maniatis, Molecular Cloning Second edition, Cold Spring Harbor Laboratory Press 1989）記載の方法によりmRNAを抽出した。次いで、オリゴキャップリンカー（配列番号：1）、オリゴdTアダプタープライマー（配列番号：2）を用いて文献（鈴木・菅野, 蛋白質 核酸 酵素, 41, 603-607 (1996). p606）の記載に従い、BAP処理、TAP処理、RNAライゲーション、第一鎖cDNAの合成とRNA

の除去を行った。次いで、PCRにより2本鎖cDNAに変換し、SfiI切断した。次いで、DraIIIで切断したベクターpME18SCG、pMFL等（菅野・丸山，蛋白質 核酸 酵素，38，472-481（1993）.p480）にcDNAの方向性を決めてクローニングした。これにより得たDNAについてDNAシーケンシング試薬（Dye Terminator Cycle Sequencing FS Ready Reaction Kit，PE Applied Biosystems）を用いてマニュアルにしたがってシーケンシング反応後、DNAシーケンサー（ABI PRISM 377，PE Applied Biosystems）でDNA塩基配列を解析した。各クローンの5'末より、それぞれ1回のみDNA配列を解析した。

【0021】

各クローンのDNA配列について、開発したcDNAの翻訳開始コドン予測プログラム（ATGpr）で翻訳開始コドンが存在するか否かにつき解析した。この解析プログラムでは数値が高いほど翻訳開始コドンである確率が高いことを示す。ただし、最高値は0.94である。

【0022】

（1）オリゴキャップ法でcDNAを作成したcDNAのうち、すでにオープンリーディングフレームがデータベースで既知のクローンについての翻訳開始コドンの解析

解析を行ったクローンのうち、決定した配列中に翻訳開始コドンが存在することがデータベースで既知のクローン（F-NT2RP1000020、F-NT2RP1000025、F-NT2RP1000039、F-NT2RP1000046）についての解析結果を以下に示す（表1）。F-NT2RP1000020(880 bp)は、88-690位が「Human neuron-specific gamma-2 enolase」（GenBank accession No.M22349）に対し相同性が96%、F-NT2RP1000025(645 bp)は、29-641位が「Human alpha-tubulin mRNA」（GenBank accession No.K00558）に対し相同性が97%、F-NT2RP1000039(820 bp)は、12-820位が「Human mRNA for elongation factor 1 alpha subunit(EF-1 alpha）」（GenBank accession No. X03558）に対し相同性が96%、F-NT2RP1000046(788 bp)は3-788位が「Human M2-type pyruvate kinase mRNA」（GenBank accession No.M23725）に対し相同性が97%である。なお、これらクローンの5'末端側の塩基配列をそれぞれ配列番号：3、4、5、6に示す。

【0023】

【表1】

F-NT2RP1000020			F-NT2RP1000025		F-NT2RP1000039		F-NT2RP1000046	
ATG No.	ATGの位置	ATGprスコア	ATGの位置	ATGprスコア	ATGの位置	ATGprスコア	ATGの位置	ATGprスコア
1	1	0.05	96	<0.94>	65	<0.90>	111	<0.94>
2	162	<0.84>	148	0.13	154	0.05	174	0.82
3	292	0.05	193	0.05	209	0.11	198	0.19
4	313	0.05	201	0.09	231	0.05	300	0.16
5	441	0.05	232	0.05	321	0.05	315	0.11

(注1) < > : 翻訳開始コドン

(注2) ATGの位置 : 5'-末よりのDNA配列でのATGの存在する塩基No.を示す。

ATG No. : 5'-末よりのDNA配列でのATGのNo.を示す。

表1が示すように、オリゴキャップ法でcDNAを作成したcDNAのうち、すでにオープンリーディングフレームがデータベースで既知で翻訳開始コドンをもつ完全長のクローンについて開発したcDNAの翻訳開始コドン予測プログラム(ATGpr)で解析した結果、翻訳開始コドンを間違いなく同定できた(データベース上の翻訳開始コドンと一致した)。

【0024】

(2) オリゴキャップ法でcDNAを作成したcDNAのうち、翻訳開始コドンが存在しないことがデータベースで既知のクローンについての翻訳開始コドンの解析解析を行ったクローンのうち、決定した配列中に翻訳開始コドンが存在しないことがデータベースで既知のクローン(F-NT2RP1000013、F-NT2RP1000054、F-NT2RP1000122)についての解析結果を以下に示す(表2)。F-NT2RP1000013(608 bp)は1-606位が「Human nuclear matrix protein 55 (nmt55) mRNA」(GenBank accession No. U89867)に対し相同性が97%、F-NT2RP1000054(869 bp)は1-86

9位が「Human signal recognition particle (SRP54) mRNA」(GenBank accession No. U51920)に対し相同性が96%、F-NT2RP1000122(813 bp)は1-813位が「H.sapiens mRNA for 2-5A binding protein」(GenBank accession No.X76388)に対し相同性が98%であった。なお、これらクローンの5'末端側の塩基配列をそれぞれ配列番号：7、8、9に示す。

【0025】

【表2】

F-NT2RP1000013			F-NT2RP1000054		F-NT2RP1000122	
ATG No.	ATGの 位置	ATGpr スコア	ATGの 位置	ATGpr スコア	ATGの 位置	ATGpr スコア
1	21	0.05	31	0.12	23	0.07
2	27	0.05	60	0.20	100	0.05
3	32	0.32	87	0.05	166	0.05
4	56	0.11	97	0.05	235	0.06
5	119	0.10	146	0.05	316	0.05
6	125	0.08	172	0.05	346	0.05
7	141	0.05	180	0.11	406	0.05
8	155	0.06	218	0.07	431	0.05
9	161	0.06	272	0.05	469	0.06
10	176	0.08	319	0.07	546	0.12
11	203	0.07	346	0.05	553	0.05
12	290	0.20	363	0.07	574	0.05
13	311	0.16	409	0.05		
14	314	0.12	480	0.07		

表2が示すようにオリゴキャップ法でcDNAを作成したcDNAのうち、すでにオープンリーディングフレームがデータベースで既知で翻訳開始コドンをもたない不

完全長のクローンについて開発したcDNAの翻訳開始コドン予測プログラム（ATGpr）で解析した結果、翻訳開始コドンを間違えて同定することはなかった。

【0026】

（3）オリゴキャップ法でcDNAを作成したcDNAのうち、新規配列のクローンについての翻訳開始コドンの解析

解析を行ったクローンのうち、翻訳開始コドンが存在すると予測された新規クローン（F-ZRV6C1000408、F-ZRV6C1000454、F-ZRV6C1000466、F-ZRV6C1000615、F-ZRV6C1000670）についての解析結果を以下に示す（表3）。なお、これらクローンの5'末端側の塩基配列をそれぞれ配列番号：10、11、12、13、14に示す。

【0027】

【表3】

F-ZRV6C1000408			F-ZRV6C1000454		F-ZRV6C1000466	
ATG	ATGの	ATGpr	ATGの	ATGpr	ATGの	ATGpr
No.	位置	スコア	位置	スコア	位置	スコア
1	85	<0.94>	5	0.05	162	<0.86>
2	208	0.22	107	<0.87>	182	0.05
3	386	0.05	153	0.05	207	0.08
4	518	0.11	201	0.08	244	0.05
5	545	0.05	211	0.05	262	0.05
6			236	0.07	303	0.11

F-ZRV6C1000615			F-ZRV6C1000670	
ATG	ATGの	ATGpr	ATGの	ATGpr
No.	位置	スコア	位置	スコア
1	85	<0.94>	120	<0.94>
2	208	0.26	187	0.54
3	386	0.05	312	0.06
4	518	0.09	388	0.05
5	545	0.05	445	0.05

(注) < > : 翻訳開始コドンと予測される

表3が示すようにF-ZRV6C1000408、F-ZRV6C1000454、F-ZRV6C1000466、F-ZRV6C1000615、F-ZRV6C1000670については、それぞれ85位、107位、162位、85位、120位の塩基「A」から始まる「ATG」が開始コドンであると判定された。このためこれらクローンは完全長cDNAクローンであると判断した。

【0028】

また、解析を行ったクローンのうち、翻訳開始コドンが存在しないと予測された新規クローン（F-ZRV6C1001410、F-ZRV6C1001197、F-ZRV6C1001472）についての解析結果を以下に示す（表4）。なお、これらクローンの5'末端側の塩基配列をそれぞれ配列番号：15、16、17に示す。

【0029】

【表4】

F-ZRV6C1001410			F-ZRV6C1001197		F-ZRV6C1001472	
ATG	ATGの	ATGpr	ATGの	ATGpr	ATGの	ATGpr
No.	位置	スコア	位置	スコア	位置	スコア
1	23	0.05	5	0.24	77	0.25
2	31	0.07	141	0.25	126	0.05
3	71	0.06	202	0.05	149	0.05
4	178	0.05	219	0.05	194	0.05
5	214	0.05	228	0.05	213	0.22
6					249	0.05
7					338	0.09
8					344	0.05
9					351	0.05
10					365	0.05

表4が示すようにF-ZRV6C1001410、F-ZRV6C1001197、F-ZRV6C1001472については、開始コドンが存在しないと判定された。このためこれらクローンは不完全長クローンであると判断した。

【0030】

【発明の効果】

本発明により、効率的に完全長のcDNAクローンを選択する方法が提供された。

本発明の方法により選択されたクローンは、完全なタンパク質を発現させることが可能である。このため本発明により、単離した遺伝子の機能解析を効率的に行うことが可能となった。

【0031】

【配列表】

- (1) 出願人の氏名又は名称： 株式会社ヘリックス研究所
- (2) 発明の名称： 完全長cDNAクローンの選択方法
- (3) 整理番号： H1-806
- (4) 出願番号：
- (5) 出願日： 1997年10月22日
- (6) 配列の数： 17

配列番号：1

配列の長さ：30

配列の型：核酸

鎖の数：一本鎖

トポロジー：直鎖状

配列の種類：他の核酸 合成RNA

配列

AGCAUCGAGU CGGCCUUGUU GGCCUACUGG

30

配列番号：2

配列の長さ：44

配列の型：核酸

鎖の数：一本鎖

トポロジー：直鎖状

配列の種類：他の核酸 合成DNA

配列

GCGGCTCAAG ACACGGCCTA TGTGGCCTTT TTTTTTTTTT TTTT

44

配列番号：3

配列の長さ：880

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

配列

```

ATGCGCCCGC GCGGCCCTAT AGGCGCCTCC TCCGCCCGCC GCCCGGGAGC CGCAGCCGCC 60
GCCGCCACTG CCACTCCCGC TCTCTCAGCG CCGCCGTCGC CACCGCCACC GCCACTGCCA 120
CTACCACCGT CTGAGTCTGC AGTCCCGAGA TCCCAGCCAT CATGTCCATA GAGAAGATCT 180
GGGCCCCGGA GATCCTGGAC TCCCGCGGGA ACCCCACAGT GGAGGTGGAT CTCTATACTG 240
CCAAAGGTCC TTTCCGGGCT GCAGTGCCCA GTGGAGCCTC TACGGGCATC TATGAGGCCC 300
TGGAGCTGAG GGATGGAGAC AAACAGCGTT ACTTAGGCAA AGGTGTCCTG AAGGCAGTGG 360
ACCACATCAA CTCCACCATC GCGCCAGCCC TCATCAGCTC AGGTCTCTCT GTGGTGGAGC 420
AAGAGAAACT GGACAACCTG ATGCTGGAGT TGGATGGGAC TGAGAACAAA TCCAAGTTTG 480
GGGCCAATCC ATCCTGGGTG TGTCTCTGGC CGTGTGTAAG GCANGGGCAA CTGAACNGGA 540
ACTGCCCCTG TATCGCCACA TTGCTCAGCT TGGNCGGGAA CTCANACCTC ATCCTGCCTG 600
TTGCCGGCCT TCAACGTGAT CAATGGTTGG CTTCTCATGC CTGGCAACAA ANCTGGCCAT 660
TGCNGGAATT TTCATGATCC TCCCNTTGG GAAACTGAAA AACTTTCCGG AATGCCCNTC 720
CAACTAAGTT GCAAAAGGTC TACCNATACC CCCCAGGGG AATTCCTCCA AGGGAACAAA 780
TNCCCGGGAA AGGAATGCCC CCCAATTNTT NGGGGGAATA AAAGGTGGGC TTTGCCCCCC 840
CATTTTCCTG GAAAAACNA TNAAAACCCT TGGGAAACTT 880

```

配列番号：4

配列の長さ：645

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

配列

```

TGTGCGTTAC TTACCTCNAC TCTTAGCTTG TCGGGGACGG TAACCGGGAC CCGGTGTCTG 60

```

CTCCTGTCGC CTTCGCCTCC TAATCCCTAG CCACTATGCG TGAGTGCATC TCCATCCACG 120
 TTGGCCAGGC TGGTGTCCAN ATTGGCAATG CCTGCTGGGA GCTCTACTGC CTGGAACACG 180
 GCATCCAGCC CGATGGCCAG ATGCCAAGTG ACAAGACCAT TGGGGGAGGA GATGACTCCT 240
 TCAACACCTT CTTCAGTGAG ACGGGCGCTG GCAANCACGT GCCCCGGGCT GTGTTTGTAG 300
 ACTTGGAACC CACAGTCATT GATGAAGTTC GCACTGGCAC CTACCGCCAG CTCTTCCACC 360
 CTGAGCAGCT CATCNCAGGC AAGGAAGATG CTGCCAATAA CTATGCCCGA GGGCACTACA 420
 CCATTGGCAA GGAGATCATT GACCTTGTGT TGGACCGAAT TCGCAAGCTG GCTGACCANT 480
 GCACCGGTCT TCANGGCTTC TTGGTTTTC ACAGCTTTGG TGGGGGAACT GGTTCCTGGGT 540
 TCACCTCCCT GCTCATGGAA CGTCTCTCAG TTGATTATGG CAAGAAATCC AAGCTGGAGT 600
 TCTCCATTTA CCCAGCACCC CNGGTTTCCN CNGCTGTANT TNGAA 645

配列番号 : 5

配列の長さ : 820

配列の型 : 核酸

鎖の数 : 二本鎖

トポロジー : 直鎖状

配列の種類 : cDNA to mRNA

配列

CTTTTTTCGC AACGGGTTTG CCGCCAGAAC ACAGGTGTCG TGAAAACTAC CCCTAAAAGC 60
 CAAAATGGGA AAGGAAAAGA CTCATATCAA CATTGTCGTC ATTGGACACG TAGATTCGGG 120
 CAAGTCCACC ACTACTGGCC ATCTGATCTA TAAATGCGGT GGCATCGACA AAAGAACCAT 180
 TGAAAAATTT GAGAAGGAGG CTGCTGAGAT GGGAAAGGGC TCCTTCAAGT ATGCCTGGGT 240
 CTTGGATAAA CTGAAAGCTG AGCGTGAACG TGGTATCACC ATTGATATCT CCTTGTGGAA 300
 ATTTGAGACC AGCAAGTACT ATGTGACTAT CATTGATGCC CCAGGACACA GAGACTTTAT 360
 CAAAAACATG ATTACAGGGA CATCTCAGGC TGACTGTGCT GTCCTGATTG TTGCTGCTGG 420
 TGTTGGTGAA TTTGAAGCTG GTATCTCCAA GAATGGGCAG ACCCGAGAGC ATGCCCTTCT 480
 GGCTTACACA CTGGGTGTGA AACAACTAAT TGTCGGTGTT AACAAAAATGG ATTCACTGAN 540
 CCACCCTACA GCCAGAAGAA ATATGANGAA ATTGTAAAGG AAGTCAGCAC TTACATTAAG 600
 AAAATTGGCT ACAACCCCGA CACAGTANCA TTTGTGCCAA TTTCTGGTTG GAATGGTGAC 660

AACATGCTGG AACCAANTGC TAACATGCCT TGGTTCCAGG GATGGAAAAT CCCCCNTTAA 720
 GGATGGCNAT GCCATTGGAA CCCCCCTGCT TGAAGGCTCT GGANTGCATC CTANCACCAA 780
 CTCCTTCAAA TTGAAAAACC CCTTGCNCCC GCCTCCNCCA 840

配列番号 : 6

配列の長さ : 788

配列の型 : 核酸

鎖の数 : 二本鎖

トポロジー : 直鎖状

配列の種類 : cDNA to mRNA

配列

GAGGCTGAGG CAGTGGCTCC TTGCACAGCA GCTGCACGCG CCGTGGCTCC GGATCTCTTC 60
 GTCTTTGCAG CGTAGCCCGA GTCGGTCAGC GCCGGAGGAC CTCAGCAGCC ATGTGGAAGC 120
 CCCATAGTGA AGCCGGGACT GCCTTCATTC AGACCCAGCA GCTGCACGCA GCCATGGCTG 180
 ACACATTCCT GGAGCACATG TGCCGCCTGG ACATTGATTC ACCACCCATC ACAGCCCGGA 240
 ACACTGGCAT CATCTGTACC ATTGGCCCAG CTTCCTCGATC AGTGGAGACG TTGAAGGAGA 300
 TGATTAAGTC TGGAATGAAT GTGGCTCGTC TGAACCTCTC TCATGGAACCT CATGAGTACC 360
 ATGCGGAGAC CATCAAGAAT GTGCGCACAG CCACGGAAAG CTTTGCTTCT GACCCCATCC 420
 TCTACCGGCC CGTTGCTGTG GCTCTAGACA CTAAAGGACC TGAGATCCGA ACTGGGCTCA 480
 TCAAGGGCAG CGGCACTGCA GAGGTGGAGC TGAAGAATGG AGCCACTCTC AAAATCACGC 540
 TGGATAATGC CTACATGGAA AAGTGTGACG AGAACATCCT GTGGCTGGAC TACAAGAACA 600
 TCTGCAAGGT GGTGGAAGTG GGCAACAAGA TCTACGTGGA TGATGGGCTN ATTTCTCTCC 660
 AGGTGAACAC AAAGGTGCCG ACTTCCTGGG TGACNGANGT GGAAAATGGT GGCTCCTTGG 720
 GCNCAAGAAA GGTGTGAACT TCCTGGGGCT GCTGTGGANT TGCCTGCTGT GTCNGAAAAA 780
 GACATCCA 788

配列番号 : 7

配列の長さ : 608

配列の型 : 核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

配列

```

ACAGCCTGGC TCCTTTGAGT ATGAATATGC CATGCGCTGG AAGGCACTCA TTGAGATGGA 60
GAAGCAGCAG CAGGACCAAG TGGACCGCAA CATCNAGGAG GCTCGTGAGA AGCTGGAGAT 120
GGAGATGGAA GCTGCACGCC ATGAGCACCA GGTCATGCTA ATGAGACAGG ATTTGATGAG 180
GCGCCAAGAA GAACTTCGGA GGATGGAAGA GCTGCACAAC CAAGANGTGC AAAAACGAAA 240
GCAACTGGAG CTCAGGCAGG AGGAANAGCG CAGGCGCCGT GAAGAANAGA TGCGGCGGCA 300
GCAAGAAGAA ATGATGCGGC GACNGCAGGA AGGATTCAAG GGAACCTTCC CTGATGCGAG 360
AGAGCAGGAG ATTCGGATGG GTCNGATGGC TATGGGAGGT GCTATGGGCA TAAACNACAG 420
ATGTGCCATG CCCCCTGCTC CTGTGCCAGC TGGTACCCCA GCTCCTCCAG GACCTGCCAC 480
TATTATGCCG GATGGAACTT TGGGATTGAC CCCACCNACA ACTGAACGCT TTGGTCNGGC 540
TGCTACNATG GAANGAATTG GGGCAATTGG TGGAATCCTT CCTGCATTCTN ACCGTGCAGC 600
TCCTGGGA 608
    
```

配列番号：8

配列の長さ：869

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

配列

```

ATATTAAACT AGTGAAGCAA CTAAGAGAAA ATGTTAAGTC TGCTATTGAT CTTGAAGAGA 60
TGGCATCTGG TCTTAACAAA AGAAAAATGA TTCAGCATGC TGTATTTAAA GAACTTGTGA 120
AGCTTGTAGA CCCTGGAGTT AAGGCATGGA CACCCACTAA AGGAAAACAA AATGTGATTA 180
TGTTTGTTGG ATTGCAAGGG AGTGGTAAAA CAACAACATG TTCAAAGCTA GCATATTATT 240
ACCAGAGGAA AGGTTGGAAG ACCTGTTTAA TATGTGCAGA CACATTCAGA GCAGGGGCTT 300
TTGACCAACT AAAACAGAAT GCTACCAAAG CAAGAATTCC ATTTTATGGA AGCTATACAG 360
    
```



```

AAATGGATCC TGTCATCATT GCTTCTGAAG GAGTAGAGAA ATTTAAAAAT GAAAATTTTG 420
AAATTATTAT TGTTGATACA AGTGGCCGCC ACAAACAAGA AGACTCTTTG TTTGAAGAAA 480
TGCTTCAAGT TGCTAATGCT ATACAACCTG ATAACATTGT TTATGTGATG GATGCCTCCA 540
TTGGGCAGGC TTGTGAAGCC CAGGCTAAGG CTTTTAAAGA TAAAGTAGAT GTACCTCAGT 600
AATAGTGACA AAACCTGATG GCCATGCAAA ANGAAGTGGT GCACTCAGTG CAGTCGCTGC 660
CACAAAAAAT CCGATTATTT TCATTGGTAC AGGGGGAACA TATANATGAC TTTGAACCTT 720
TCAAAAACAC AGCCTTTTAT TAACAACTT CTTGGTATNG GCGACATTGA AAGGACTGAT 780
AAATAAAGTC CACNAATTGA AATTTGGATG ACNATGNAAA CCCTTATTGA AAAAATTGAA 840
ACATNGTCCA GTTTTACTTT GCGAAACNT 869

```

配列番号 : 9

配列の長さ : 813

配列の型 : 核酸

鎖の数 : 二本鎖

トポロジー : 直鎖状

配列の種類 : cDNA to mRNA

配列

```

GTTGTGGTAT CTGTATTAAG AAATGCCCCCT TTGGCGCCTT ATCAATTGTC AATCTACCAA 60
GCAACTTGGA AAAAGAAACC ACACATCGAT ATTGTGCCAA TGCCTTCAAA CTTACAGGT 120
TGCCTATCCC TCGTCCAGGT GAAGTTTGG GATTAGTTGG AACTAATGGT ATTGGAAAGT 180
CAACTGCTTT AAAAATTTTA GCAGGAAAAC AAAAGCCAAA CCTTGGAAG TACGATGATC 240
CTCCTGACTG GCAGGAGATT TTGACTTATT TCCGTGGATC TGAATTACAA AATTACTTTA 300
CAAAGATTCT AGAAGATGAC CTAAAAGCCA TCATCAAACC TCAATATGTA GACCAGATTC 360
CTAAGGCTGC AAAGGGGACA GTGGGATCTA TTTTGGACCG AAAAGATGAA ACAAAGACAC 420
AGGCAATTGT ATGTCAGCAG CTTGATTTAA CCCACCTAAA AGAACGAAAT GTTGAAGATC 480
TTTCAGGAGG AGAGTTGCAG AGATTTGCTT GTGCTGTCGT TTGCATACAG AAAGCTGATA 540
TTTTCATGTT TGATGAGCCT TCTAGTTACC TAGATGTCAA GCAGCGTTTA AAGGCTGCTA 600
TACTATACG ATCTCTAATA AATCCAGATA GATATATCAT TGTGGTGGAA CATGATCTAA 660
GTGTATTAGA CTATCTCTCC GACTTCATCT GCTGTTTATA TGGTGTACCA AGCGCCTATG 720

```

GAATTGTCAC TATGCCTTTT AGTGTTAGAA AAGGCATAAA CNTTTTTTGG ATGGGTATGT 780
TCCAACAGAA AACTTGANAA TCNNAAATGC NTC 813

配列番号 : 10

配列の長さ : 655

配列の型 : 核酸

鎖の数 : 二本鎖

トポロジー : 直鎖状

配列の種類 : cDNA to mRNA

配列

AGACTCTCAC CGCAGCGGCC AGGAACGCCA GCCGTTACG CGTTCGGTCC TCCTTGGCTG 60
ACTCACCGCC CTCGCCGCCG CACCATGGAC GCCCCAGGC AGGTGGTCAA CTTTGGGCCT 120
GGTCCCGCCA AGCTGCCGCA CTCAGTGTTG TTAGAGATAC AAAAGGAATT ATTAGACTAC 180
AAAGGAGTTG GCATTAGTGT TCTTGAAATG AGTCACAGGT CATCAGATTT TGCCAAGATT 240
ATTAACAATA CAGAGAATCT TGTGCGGGAA TTGCTAGCTG TTCCAGACAA CTATAAGGTG 300
ATTTTTCTGC AAGGAGGTGG GTGCGGCCAG TTCAGTGCTG TCCCCTTAAA CCTCATTGGC 360
TTGAAAGCAG GAAGGTGTGC GGACTATGTG GTGACAGGAG CTTGGTCAGC TAAGGCCGCA 420
GAAGAAGCCA AGAAGTTTGG GACTATAAAT ATCGTTCACC CTAAACTTGG GAGTTATACA 480
AAAAATTCCAG ATCCAAGCAC CTGGAACCTC AACCCANATG CCTCCTACGT GTTTTATTGC 540
NCAAATGAAA CGGTGCATGG TGTTGANTTT GACTTTATAC CCNATGTCAA GGGAACANTA 600
CTGGTTTGTG ACATTTTCCT CCAACTTCCT GTCCAANCCA ATTGNATGTT TCCAA 655

配列番号 : 11

配列の長さ : 599

配列の型 : 核酸

鎖の数 : 二本鎖

トポロジー : 直鎖状

配列の種類 : cDNA to mRNA

配列

AAAGATGCGC AGGCGCCGTG TGGCACTCGG CGGTCGAAAG GGGAGTTCAA GGAGACGGGG 60
 GCGACGCGGC TGAGGGCTTC TCGTCGGGGT CGGGGCTGCA GCCGTCATGC CGGGGATAGT 120
 GGAGCTGCCC ACTCTAGAGG AGCTGAAAAGT AGATGAGGTG AAAATTAGTT CTGCTGTGCT 180
 TAAAGCTGCG GCCCATCACT ATGGAGCTCA ATGTGATAAG CCCAACAAGG AATTTATGCT 240
 CTGCCGCTGG GAANAGAAAG ATCCGAGGCG GTGCTTAGAG GAAGGCAAAC TGGTCAACAA 300
 GTGTGCTTTG GACTTCTTTA GGCAGATAAA ACGTCACTGT GCAGAGCCTT TTACAGAATA 360
 TTGGACTTGC ATTGATTATA CTGGCCAGCA GTTATTTCTG CACTGTCGCA AACAGCAGGC 420
 AAAGTTTGAC NAGTGTGTGC TGGACAAACT GGGCTGGGTG CGGCCTGACC TGGGAAAAC 480
 GTCAAAGGTC ACCAAAGTGA AAACAGATCN ACCTTTACCG GANAATCCCT ATCACTCAAG 540
 AACAAGAACG GATCCCAGCC CTGANATCNA AGGAAATCTG CANCCTGCCA CACATGGCA 599

配列番号 : 12

配列の長さ : 597

配列の型 : 核酸

鎖の数 : 二本鎖

トポロジー : 直鎖状

配列の種類 : cDNA to mRNA

配列

ATATCCGGAG TAGACGGAGC CGCAGTAGAC GGATCCGCGG CTGCACCAAA CACTGCCCCCT 60
 CGGAGCCTGG TAGTGGGCCA CAAGCCCCCA GTCCCAGAGG CGTGATTTTC TGGCATCCTT 120
 AAATCTTG TG TCAAGGATTG GTTATAATAT AACCAGAAAC CATGACGGCG GCTGAGAACG 180
 TATGCTACAC GTTAATTAAC GTGCCAATGG ATTCAGAACC ACCATCTGAA ATTAGCTTAA 240
 AAAATGATCT AGAAAAAGGA GATGTAAAGT CAAAGACTGA AGCTTTGAAG AAAGTAATCA 300
 TTATGATTCT GAATGGTGAA AACTTCCTG GACTTCTGAT GACCATCATT CGTTTTGTGC 360
 TACCTCTTCA GGATCACACT ATCAAGAAAT TACTTCTGGT ATTTTGGGAG ATTGTTCTTA 420
 AAACAACCTC AGATGGGAGA CTTTACATG AGATGATCCT TGTATGTGAT GCATACAGAA 480
 AGGATCTTCA ACATCCTAAT GAATTTATTC NAAGGATCTA CTCTTCGTTT TCTTTGCAAA 540
 TTGAAANAAA CANAATTGCT AAAACCTTTA ATGCCANCTA TNCCTGCATT TTTGGGA 599

配列番号 : 13

配列の長さ : 634

配列の型 : 核酸

鎖の数 : 二本鎖

トポロジー : 直鎖状

配列の種類 : cDNA to mRNA

配列

```

AGACTCTCAC CGCAGCGGCC AGGAACGCCA GCCGTTACAG CGTTCGGTCC TCCTTGGCTG   60
ACTCACC GCC CTGCGCGCCG CACCATGGAC GCCCCAGGC AGGTGGTCAA CTTTGGGCCT  120
GGTCCCGCCA AGCTGCCGCA CTCAGTGTTG TTAGAGATAC AAAAGGAATT ATTAGACTAC  180
AAAGGANTTG GCATTAGTGT TCTTGAAATG AGTCACAGGT CATCAGATT TGCCAAGATT  240
ATTAACAATA CAGAGAATCT TGTGCGGGAA TTGCTAGCTG TTCCAGACAA CTATAAGGTG  300
ATTTTCTGCG AAGGAGGTGG GTGCGGCCAG TTCAGTGCTG TCCCCTTAAA CCTCATTGGC  360
TTGAAAGCAG GAANGTGTGC GGAATATGTG GTGACAGGAG CTTGGTCAGC TAAGGCCGCA  420
NAANAAGCCA AGAANTTTGG GACTATAAAT ATCGTTCACC CTAAACTTGG GAGTTATACA  480
AAAATTCCAG ATCCAAGCAC CTGGAACCTC AACCAGATG CCTCCTACGT GTATTATTGC  540
GCNAATGAAA CNGTGCATGG TGTGGANTCT GACTTTATAC CCGATGTCNA GGGAACATAC  600
TGGTTTGTGA CATGTCCTCA AACTTCCCGT CCNA                                634
    
```

配列番号 : 14

配列の長さ : 757

配列の型 : 核酸

鎖の数 : 二本鎖

トポロジー : 直鎖状

配列の種類 : cDNA to mRNA

配列

```

AGTCTGCGGT GGGCTANCGG ACGGTCCGGC TTCCGGCGGC CGTTTCTGTC TCTTGCTGGC   60
TGTCTCGCTG AATCGCGGCC GCCTTCTCAT CGCTCCTGGA AGGTCCCGAG CGCGACACCA  120
TGTCGGAACC CGGGGCGGCC GGCGGCGAAG ACNGCTCGGC CGGATTGGAA GTGTCGGGCC  180
    
```

TGCANAATGT GCGGACGTG TCGGTGCTGC ANAAGCACCT GCGCAAGCTG GTGCCGCTGC 240
 TGCTGGAGGA CGGCGGCGAA GCGCCGGCCG CGCTGGAGGC GCGCTGGAG GAGAAGAGCG 300
 CCCTGGAGCA GATGCGCAAG TTCCTTTCGG ACCCGCACGT CCACACGGTG CTGGTGGAGC 360
 GCTCCACGCT CAAAGTGGAC GTCGGTGATG AAGGAGAAGA AGAAAAAGAA TTCATTTCTT 420
 ATAACATCAA CNTAGACATT CACTATGGGG TTAAATCCAA TAGCTTGGCA TTCATTAAAC 480
 GTACTCCCGT GATTGATGCA GATAAACCCG TGTCTTCTCA NCTCCGGGTC CTTACACTCA 540
 GTGAANACTC NCCCTACNAA AACTTTGTCAT TCTTTCATTA ACAATGCAGT GGCTCCTTTT 600
 TTTAANTCCT ACATTAAAAA ATCTGGCAAG GCAAACAGGG ATGGTGATAA AATGGCTCCT 660
 TCCNTTGAAA AAAAAATTGC CGAACTCNAA ATNGGACTCC TTCCCTTGCA NCAAAATTTT 720
 TGAAATTCCG GAAAATCANC CTGCCCAATT CCTCCCC 757

配列番号 : 15

配列の長さ : 300

配列の型 : 核酸

鎖の数 : 二本鎖

トポロジー : 直鎖状

配列の種類 : cDNA to mRNA

配列

ATCATTTCCT TATTTATATT TCATGTTGGA ATGCTTAAAT CGATAACCTT TGTATTTTGA 60
 AGTGCGCGAC ATGGAAGGTG ATCTGCAAGA GCTGCATCAG TCAAACACCG GGGGATAAAT 120
 CTGGATTGG GTTCCGGCGT CAAGGTGAAG ATAATACCTA AAGAGGAACA CTGTAAAATG 180
 CCAGAAGCAG GTGAANAGCA ACCACAAGTT TAAATGAAGA CAAGCTGAAA CAACGCAAGC 240
 TGGTTTTATA TTAGATATTT GACTTAAACT ATCTCAATAA AGTTTTGCAG CTTTCACCAC 300

配列番号 : 16

配列の長さ : 313

配列の型 : 核酸

鎖の数 : 二本鎖

トポロジー : 直鎖状

配列の種類 : cDNA to mRNA

配列

```

AAAGATGGCG GCGGGGGAGG TAGGCAGAGC AGGACGCCGC TGCTGCCGCC GCCACCGCCG 60
CCTCCGCTCC AGTCGCCTCC GGTCCCTCAA ACTCACACCT CCCGGGAGGA GCTGTCCTGG 120
CGCCGGGTCC CGCGGGGAAA ATGGTGGAGC CAGGGCAAGA TTTACTGCTT GCTGCTTTGA 180
GTGAGAGTGG AATTAGTCCG AATGACTCTT TGATATTGAT GGTGGAGATG CANGGCTTGC 240
AACTCCAATG CCTACCCCGT CAGTTCAGCA NTCAGTGCCA CTTANTGCAT TANAACANG 300
TTTGGAGACC GAA 313

```

配列番号 : 17

配列の長さ : 667

配列の型 : 核酸

鎖の数 : 二本鎖

トポロジー : 直鎖状

配列の種類 : cDNA to mRNA

配列

```

ACTGCCGGGC TCGGCGTGAG TCGCTGCGGG GCTGACGGGG TGGCAGTGCG GCGGGTTACG 60
GCCTGGTCAG ACCATAATGA CTTCAGCAAA TAAAGCAATC GAATTACAAC TACAAGTGAA 120
ACAAAATGCA GAAGAATTAC AAGACTTTAT GCGGGATTTA GAAAACTGGG AAAAAGACAT 180
TAAACAAAAG GATATGGAAC TAAGAAGACA GAATGGTGTT CCTGAAGAGA ATTTACCTCC 240
TATTCGAAAT GGAATTTTA GGAAAAAGAA GAAAGGCAAA GCTAAAGAGT CTTCCCCAAA 300
ACCANAGAGG AAAACACNAA AAACAGGATA AAATCTTATG ATTATGANGC ATGGGCAAAA 360
CTTGATGTGG ACCGTATCCT TGATGAGCTT GACAAAGACG ATAGTACCCA TGAGTCTCTG 420
TCTCAAGAAT CAGAGTCGGA AGAAGATGGG ATTCATGTTG ATTCNCNAAA GGCTCTTGTT 480
TAAAAAGAAA AGGGCNATAA AACTTCCAC AAGGAAAATA TGATGAAGCA ATTGACTGCT 540
ACACNAAAGG CNTGGATGCC GATCCATATN ATCCCGTGTT GCCAACGAAC ANAACNTCCG 600
CATATTTTAG ACTGAAAAAA TTTGCTGTTG CTGAATCTGA TTGTTATTAN CANTTGCCT 660
TGAAATA 667

```

【書類名】 要約書

【要約】

【課題】 効率的に完全長のcDNAクローンを選択する方法、および完全長率の高いcDNAライブラリーの作製方法を提供することを課題とする。

【解決手段】 完全長率の高いcDNAライブラリー作成法で作製したcDNAライブラリーに含まれるクローンの5'領域の塩基配列を決定し、独自に開発したcDNAの翻訳開始コドン予測ソフトを利用してこの5'領域に翻訳開始ATGが存在するか否か、およびその存在位置の検討を行った。その結果、独自に開発したソフトが翻訳開始コドンの存在の有無および存在位置を的確に判定することができ、このソフトにより翻訳開始コドンが存在すると判定されたクローンをcDNAライブラリーから選択することにより効率的に完全長のcDNAクローンを選択することが可能であることを見いだした。さらに、選択したクローンを混合することにより、完全長率の極めて高いcDNAライブラリーを作製できることを見いだした。

【選択図】 なし

【書類名】 職権訂正データ
【訂正書類】 特許願

<認定情報・付加情報>

【特許出願人】

【識別番号】 597059742

【住所又は居所】 千葉県木更津市矢那1532番地3

【氏名又は名称】 株式会社ヘリックス研究所

【代理人】 申請人

【識別番号】 100102978

【住所又は居所】 茨城県土浦市卸町1-1-1 関鉄つくばビル6階
清水国際特許事務所

【氏名又は名称】 清水 初志

【選任した代理人】

【識別番号】 100108774

【住所又は居所】 茨城県土浦市卸町1-1-1 関鉄つくばビル6階
清水国際特許事務所

【氏名又は名称】 橋本 一憲

出 願 人 履 歴 情 報

識別番号 [597059742]

1. 変更年月日	1997年 4月28日
[変更理由]	新規登録
住 所	千葉県木更津市矢那1532番地3
氏 名	株式会社ヘリックス研究所

This Page Blank (uspio)